
M O S E R

(TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Sat May 13 08:08:38 2000; Maspar time 23.60 Seconds
668,443 Million cell updates/sec
Tabular output not generated.

Title: >US-09-331-631-3
Description: (1-666) from US09331631.pep (1 of 5)
Perfect Score: 4913
Sequence: 1 MAINTSNLCSLFLSLFL.....SSRSTKQQPLVSIIDVGF 666

Scoring table:
PAM 150
Gap 11

Searched: 188963 segs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
1:geneseqp

Statistics: Mean 37.214; Variance 174.766; scale 0.213

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	4913	100.0	666	1	Macadamia integrifolia	0.00e+00
2	4746	96.6	666	1	Macadamia integrifolia	0.00e+00
3	4496	91.5	625	1	Macadamia integrifolia	0.00e+00
4	1353	27.5	566	1	Sequence encoded by 67	3.64e-107
5	1236	25.2	590	1	Gossypium hirsutum ant	1.29e-96
6	1193	22.3	525	1	Theobroma cacao antimi	9.54e-93
7	1106	22.5	593	1	Zea mays antimicrobial	6.29e-85
8	1036	21.1	489	1	G. max SBP2 protein.	1.19e-78
9	998	20.3	524	1	G. max SBP2 protein.	3.01e-75
10	931	20.0	605	1	Glycine max antimicrob	9.98e-74
11	933	19.0	444	1	G. max truncated SBP1	1.94e-69
12	933	19.0	626	1	Peanut allergen Ara hi	1.94e-69
13	922	18.8	614	1	Peanut allergen Ara hi	1.86e-68
14	922	18.8	614	1	Arachis hypogaea antim	1.36e-68
15	890	18.1	409	1	G. max truncated SBP2	1.33e-65
16	842	17.1	637	1	Hordeum vulgare anti	2.49e-61
17	180	3.7	395	1	Mouse SRY-related prot	1.10e-04
18	145	3.5	303	1	Diofilaria immitis pa	3.15e-04
19	144	3.0	186	1	Trypanosoma cruzi anti	4.48e-02
20	147	3.0	567	1	Polypeptide fragment e	3.20e-02
21	136	2.8	432	1	Human regulatory molec	1.99e-01
22	130	2.6	265	1	HTIV-1 protein express	5.28e-01
23	127	2.6	316	1	P.falciparum USA-R-NR	8.58e-01

24	127	2.6	493	1	R26944	P.falciparum USA gene	8.58e-01
25	130	2.6	541	1	W37148	Mammalian Ena (Mena).	5.28e-01
26	126	2.6	562	1	R70491	Leucocytozoan protozoa	1.01e+00
27	130	2.6	783	1	W37151	Mouse neural Mena+ pro	5.28e-01
28	130	2.6	787	1	W37152	Mouse neural Mena+ pr	5.28e-01
29	130	2.6	802	1	W37153	Mouse neural Mena+ p	5.28e-01
30	129	2.6	1529	1	R37985	CORX potassium channel	6.21e-01
31	125	2.5	1382	1	W31867	Human metastasis-assoc	1.18e+00
32	116	2.4	190	1	R82664	E. coli RNase E.	4.91e+00
33	116	2.4	351	1	W44681	Escherichia coli RNase	4.91e+00
34	120	2.4	351	1	R24393	Sequence of Histidine	2.62e+00
35	116	2.4	436	1	W03662	Human 70k U1 snRNP pro	4.91e+00
36	118	2.4	521	1	W74802	Human secreted protein	3.59e+00
37	116	2.4	614	1	R82630	70k autoantigen, part	4.91e+00
38	117	2.4	1178	1	W30763	Mannose-1-phosphate tr	4.20e+00
39	111	2.3	28	1	W62841	Stenocarpus sinuatus a	1.07e+01
40	113	2.3	167	1	W89768	Staphylococcus aureus	7.82e+00
41	112	2.3	325	1	W59645	Amino acid sequence of	9.13e+00
42	114	2.3	1311	1	W72971	Precis coenia patched	6.70e+00
43	114	2.3	1311	1	W52197	Precis coenia (butleri	6.70e+00
44	112	2.3	1343	1	W31866	Mouse metastasis-assoc	9.13e+00
45	110	2.2	482	1	Y07067	Renal cancer associate	1.24e+01

ALIGNMENTS

RESULT	ID	Score	Query Match	Length	ID	Description	Pred. No.
1	W62829	100.0%	666	666	1	Macadamia integrifolia	0.00e+00
2	W62829	96.6%	625	625	1	Macadamia integrifolia	0.00e+00
3	W62829	91.5%	625	625	1	Macadamia integrifolia	0.00e+00
4	W62829	27.5%	566	566	1	Sequence encoded by 67	3.64e-107
5	W62829	25.2%	590	590	1	Gossypium hirsutum ant	1.29e-96
6	W62829	22.3%	525	525	1	Theobroma cacao antimi	9.54e-93
7	W62829	22.5%	593	593	1	Zea mays antimicrobial	6.29e-85
8	W62829	21.1%	489	489	1	G. max SBP2 protein.	1.19e-78
9	W62829	20.3%	524	524	1	G. max SBP2 protein.	3.01e-75
10	W62829	20.0%	605	605	1	Glycine max antimicrob	9.98e-74
11	W62829	19.0%	444	444	1	G. max truncated SBP1	1.94e-69
12	W62829	19.0%	626	626	1	Peanut allergen Ara hi	1.94e-69
13	W62829	18.8%	614	614	1	Peanut allergen Ara hi	1.86e-68
14	W62829	18.8%	614	614	1	Arachis hypogaea antim	1.36e-68
15	W62829	18.1%	409	409	1	G. max truncated SBP2	1.33e-65
16	W62829	17.1%	637	637	1	Hordeum vulgare anti	2.49e-61
17	W62829	3.7%	395	395	1	Mouse SRY-related prot	1.10e-04
18	W62829	3.5%	303	303	1	Diofilaria immitis pa	3.15e-04
19	W62829	3.0%	186	186	1	Trypanosoma cruzi anti	4.48e-02
20	W62829	3.0%	567	567	1	Polypeptide fragment e	3.20e-02
21	W62829	2.8%	432	432	1	Human regulatory molec	1.99e-01
22	W62829	2.6%	265	265	1	HTIV-1 protein express	5.28e-01
23	W62829	2.6%	316	316	1	P.falciparum USA-R-NR	8.58e-01

Db	241	EDGEKQSDNYPYPERELSTFRPEEESHIVLNFGRSKLPAKLNRYLVLEAPNA	300			
Oy	241	EDGEKQSDNYPYDERLSLSTRFEEBSHIVLENTFGSKLLAALKNRYLVLEAPNA <td>3000</td>	3000			
Db	301	FVLPLHLDADAILLVYGGGALKMIHRDNRESYNLECGDIVRIPAGTFFYLINRDNREL	360			
Oy	301	FVLPLHLDADAILLVYGGGALKMIHRDNRESYNLECGDIVRIPAGTFFYLINRDNREL	360			
Db	361	HIAKLQIISIPGVCKEFPAGGONPEEYISTEKKELALNQAERLGVLGQREQV	420			
Oy	361	HIAKLQIISTPGQKKEFPAGGONPEEYISTEKKELALNQAERLGVLGQREQV	420			
Db	421	IISASQEDIRELTRDSSRRMWHIRGESSRSGYNLFNKRPLYSNKGAYEYKPEDYR	480			
Oy	421	IISASQEDIRELTRDSSRRMWHIRGESSRSGYNLFNKRPLYSNKGAYEYKPEDYR	480			
Db	481	QLOQMDVSVFIANTIQSGMGPFFENTSKTKVYVYASGEADVEMACPHLSGRHGGRGGR	540			
Oy	481	QLOQMDVSVFIANTIQSGMGPFFENTSKTKVYVYASGEADVEMACPHLSGRHGGRGGR	540			
Db	541	HEEEDVAYEYKRLSKREAIYVPVGHVYVSSGNENLLFPAFGIAONNHNENLAGR	600			
Oy	541	HEEEDVAYEYKRLSKREAIYVPVGHVYVSSGNENLLFPAFGIAONNHNENLAGR	600			
Db	601	ERNVLQOIEPQAMELAFAPRKEVEEELFNSODESIFPPGPRHQOQSSRSTKQOQPLYSI	660			
Oy	601	ERNVLQOIEPQAMELAFAPRKEVEEELFNSODESIFPPGPRHQOQSSRSTKQOQPLYSI	660			
Db	661	LDYVGF 666				
Oy	661	LDYVGF 666				
RESULT 2						
ID	W62828 standard; protein; 666 AA.					
AC	W62828.					
DT	27-OCT-1998 (first entry)					
DE	Macadamia integrifolia antimicrobial protein.					
KW	antimicrobial protein; infestation; control.					
OS	Macadamia integrifolia.					
FN	Key Location/Qualifiers					
FT	Peptide 1..28					
FT	Protein /note= "signal peptide"					
FT	29..666					
FT	Protein /note= "mature protein"					
PN	M09827805-A1.					
PD	02-JUL-1998.					
PF	22-DEC-1997; A00874.					
PR	20-DEC-1996; A0-004275.					
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.					
PI	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;					
DR	WPI: 98-377279/32.					
DR	N-PSDB: V42310.					
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -					
FT	useful for controlling microbial infestations of plants or mammals					
PS	Claim 1; Page 34-36; 96pp; English.					
CC	The sequence is that of an antimicrobial protein which can					
CC	be used to control microbial infestations in plants and mammalian					
CC	animals.					
SO	Sequence 666 AA;					
Query Match 96.6%; Score 4746; DB 1; Length 666;						
Best Local Similarity 96.1%; Pred. No. 0.00e+00;						
Matches 640; Conservative 16; Mismatches 10; Indels 0; Gaps						
Db	1	MAINTSNLCSLFLSLFSLSTVYSLASEPDRQYEYCKRQCMQLETSGQMRRCVSGCD	60			
Oy	1	MAINTSNLCSLFLSLFSLSTVYSLASEPDRQYEYCKRQCMQLETSGQMRRCVSGCD	60			
Db	61	KRFEDIDMSYDNOEDPQTCOCOCORRCROESGPRQOQYCORCKEICEEHEEYNROR	120			
Oy	61	KRFEDIDMSYDNOEDPQTCOCOCORRCROESGPRQOQYCORCKEICEEHEEYNROR	120			
Oy	61	KRFEDIDMSYDNOEDPQTCOCOCORRCROESGPRQOQYCORCKEICEEHEEYNROR	120			

Db	121	DPDOYECCKHCHORRELEPNNHMO	CGOQREERREKRRKQOQRYEEOQDEDEKTEER	180
Qy	121	DPQOQTECCOERCHNELEPHNMOT	CCOQREERREKRRKQOQRYEEOQDEDEKTEER	180
Db	181	KEEDNKRDPQOREVEDCRRRCR	EOEOPRQOQOLCRHEOQROGRGQDMNPNRGSGGR	240
Qy	181	KEEDNKRDPQOREVEDCRRRCR	EOEOPRQOQORQOORCRGQDLINPRGSGGR	240
Db	241	EEGEEGSDNRYIYDEDSLS	TRFTEEGHISYLENTYGRSKLLRALKNRYLVILLEANPNA	300
Qy	241	EEGEEKOSDNRYIYDEDSLS	TRFTEEBGHISYLENFYGRSKLLRALKNRYLVILLEANPNA	300
Db	301	FVLPTLHDADAILVYIGRGALKI	HHNDNESYNLECGDVIRIPAGTFYLINDNNERL	360
Qy	301	FVLPTLHDADAILVYIGRGALKI	HHNDNESYNLECGDVIRIPAGTFYLINDNNERL	360
Db	361	HIKFLQITSTPGQYKEFFPAGGON	PEPULSTFSKELLEALNTQTEKLYGVEGQOREG	420
Qy	361	HIKFLQITSTPGQYKEFFPAGGON	PEPULSTFSKELLEALNTQOERLGVIGQOREG	420
Db	421	IIRASOEQIRPLTRDDSESRHWH	IIRGCESSRGPYNLFNKRPILSNKYGQAYEVKREDR	480
Qy	421	IIRASOEQIRPLTRDDSESRHWH	IIRGCESSRGPYNLFNKRPILSNKYGQAYEVKREDR	480
Db	481	QLOQMDLSVFLANTQSGMMGP	FNFTSTVVVVAASEADVEMACPLLSRHRGRRGGGR	540
Qy	481	QLOQMDLSVFLANTQSGMMGP	FNFTSTVVVVAASEADVEMACPLLSRHRGRRGGGR	540
Db	541	HEEEDVHYEDVRRARLSKREAI	VLVLAGHPVVFSSGNGENLLLEAFGINQNNHENTLAGR	600
Qy	541	HEEEDVHYEDVRRARLSKREAI	VLVLAGHPVVFSSGNGENLLLEAFGINQNNHENTLAGR	600
Db	601	ERNVLOQIEPQAMELAPAPRKE	VEEESFNQODSIFFPGRQHQOQSPRSTKOQOPLVSI	660
Qy	601	ERNVLOQIEPQAMELAPAPRKE	VEEELFNSODESIEFFPGRQHQOQSSRSTKOQOPLVSI	660
Db	661	LDPVGF	666	
Qy	661	LDPVGF	666	

```

RESULT      3
ID MW82830 standard; Protein; 625 AA.
AC M62830;
DT 27-OCT-1998 (first entry)
DE Macadamia integrifolia antimicrobial protein.
OS antimicrobial protein; infestation; control.
OS Macadamia integrifolia.
FH Key
FH Peptide
FT 1.28
FT /note="signal peptide"
FT 29.666
FT Protein
FT /note="mature protein"
PN W09827805-A1.
PD 02-JUL-1998.
PE 22-DEC-1997: AU0874.
PR 20-DEC-1996: AU-004275.
PR (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PA Bower NI, Goulier KC, Green JL, Manners JM, Marcus JP;
PI WPI: 98-377279/32.
DR N-PSDB: V42316.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1, Page 43-45, 96pp, English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 625 AA;
Query Match 91.5%; Score 4496; DB 1; Length 625;
Best Local Similarity 96.6%; Pred. No. 0.00e+00;
Matches 604; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

```


[illegible]

	Query Match	24.3%	Score 1193;	DB 1;	Length 525;
	Best Local Similarity	40.6%;	Pred. No. 9,54e-93;		
	Matches	195;	Conservative 122;	Mismatches 140;	Gaps 17/;
Db	33 QERDPRQYEQCCRRPSEE-AIEEREDE-OCE--OQCERE-YKEQRROOE-E-LORO	84			
	: :				
OY	135 RHLETPRNHM-QTQQCRDRREYKKRKRQKRYEEOQAQDEDEKYIERRKKEDNRKPQORE	193			
	:				
Db	85 YQSCGRCQEEOQOGQREDOQOCQAKCMEQYKEQENG-EHENYHNHKKNRSEEGSQRPNP	143			
	: :				

QY	134	YEDCRRK-EQOE-PROOYQOORCRREDQROHGREGGLINFGQSGSGRYEEGEEKQSDNP	251
Db	144	YFPPKRSFOTFRPDEENFXILQFAENSPPLKGINDRFLAMEANPNTILPHHCDAE	203
QY	252	YFEDER-SLSTRFRREBEGHISYLENFYRSRLRLRLKRYLLVLLLEANNNARVLP7HIDAD	310
Db	204	AIYFPTNKGITTEYTHHNKESYINQCTVYVSPAGSTVYVVSODNOEKLTIYALVLPN	263
QY	311	ALLTGTGRGALKMIRHNRRESYNIECGDYIRIPAGTIFYILINDNNRHLIAEFLQ7IS	370
Db	264	SPGKVELFPFPGNNKPESYVGAFSYEYETVFNQREKLEIELEORQKNOQOQGMFR	323
QY	371	TPGQYKEFFPPAGGNPEYILTFKSKEIIEALNTOAERLRYLOQOQREGYIISASQEOIR	430
Db	334	KAKPEQIIRAISOQATSPRHRRGERLAINLTSQSPVYSNONGRFFEDACPEDESOPOMDVA	383
QY	431	ELTRDD--SESRPMWIRRGSESSRQPYLFPKRPLYSNKYGQAVEKRPEDYRQLODMVS	488
Db	384	VSARLNGALFVPHYNKATFVYVDGYIAQMACPHLSROSGQSGSGQDRDEEE	443
QY	489	VFIATITGSGNMGPENFRSTKYVVAVASGEADVEMACP7LSGR-HGGRG--GRR-HEE	544
Db	444	SEEFEGFGQOKAPLPSGDVFVAPAGHAYFFASQDOP7LAANAQNNQRI7FLAGR	503
QY	545	--EDV--HYEQVKR7LSRREIAIVPVGH7PVYFVSSGNNLL7LFAFG7INAQNNHEN7FLAGR	600
RESULT 7			
ID	W62835	standard; Protein; 593 AA.	
AC	W62835;		
DT	27-OCT-1998	(first entry)	
DE	Zea mays antimicrobial protein.		
KW	antimicrobial protein; Infestation; control.		
OS	Zea mays.		
PN	MO9827805-A1.		
PF	02-JUL-1998.		
PT	22-DEC-1997.		
RA	20-DEC-1996; AU-004275.		
PR	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.		
PI	Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;		
DR	WPI; 98-37729/32.		
PT	Novel antimicrobial protein from e.g. Macadamia integrifolia -		
PR	useful for controlling microbial infestations of plants or mammals		
PS	Claim 1, Page 58-60; 36pp; English.		
CC	The sequence is that of an antimicrobial protein which can		
CC	be used to control microbial infestations in plants and mammalian		
CC	animals.		
Q0	Sequence 593 AA;		

Query Match	22.5%	Score 1106	DB 1:	Length 593
Best Local Similarity	40.8%	Pred. No. 6.29e-85		
Matches 204	Conservative 112	Mismatches 155	Indels 29	Gaps 22
D	25	EDDNHHHGGHSGQCVRCEDR-FWHQRPRLCEÖCEEEREKROERSRHREADDRSGGS	83	
Q	182	EDDKRDPÖGQKEYEDCKRRCEÖQDPPOÖY-QÖCRRCÖQÖRÖ-QHGRGDLINPÖRG-GS	237	
D	84	SEDERÖEKÖKÖRÖRPFYEDRSEFRVRSÖEGSLVLPFDEVSLLGIRYRAVL	143	
Q	238	GRYE-EGE-EKÖSN-PFYEDERLSTRFTEBESHISYLENFYGRSLRLAKYRLAYVL	294	
D	144	EANDRSEFVDSHTDAHCICYAAEEGCVYTTIENGERSYTIKÖGHVFADAGAVTYLANT	203	
Q	295	EANNAEVLPTHLADADLLVLTGGRGLAKMIHRDNRSYMLEGCDVYRIAPAGTFYILNR	354	
D	204	DGRKLVYTKLHITISVGEFOFFGCGNRPESLFSFKSIÖRÄAYKTSDDLERLFG	263	
Q	355	DNMRLLHAKFLQYISIPGQYKEFFPAGCONPEYLSLTFSEKILEALNLTQAEHLRQVLG	414	
D	264	RHGÖDKIIVRATEDÖTRRELRHHAEGGHHPWLPFEGES-RGYPSSLDÖRPSIANÖHG	352	
Q	415	QÖ-RE-CVILIASÖQÖRRELTRODSÖS-R-RWHIRRGSGSGRPYLLFPMKRPYLSKYG	469	

D6 QLYPADRSFHDIAEHVSVSPANTAGSMASALFNTSRFKATLVYNGGVAEIVCPHQ 382
D7 | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
OY 470 QAYEVRKEDTRQLOMDVSYVFINTIITGSGMGEFFNTRSTKYVVYASGADVDMAACP-H L 528
D8 383 SQGSESERENDKRGRSREEEEESSEDEENGOGYHTIRARLSPGTAFFVYPAGHPFAVAAS 442
D9 | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
OY 529 S-GHGRG-RG-GKR-----HE--EE-EDV-H-YEQVKARLSKREAIYVPVGHVAVSS 575
D0 443 RDSNLQVCVEYHADREKEVFLGAD-NVLQTKDRPAKALSFSKAEDEVILGSRRREG 501
OY 576 GNEMLLLFAGGINANNHNENFLAGRKNVLOQLPEQAMELAFAAPRKEEELFNSDESI 635
D1 502 FLPGPEESGGHEEEQESEE 521
OY | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
OY 636 FFGPGRHQOOSSSRSTKQQ 655

RESULT 8
ID W90341 standard; protein; 489 AA.
AC W90341;
DT 24-MAY-1999 (first entry)
DE G.max SBP2 protein.
KW SBP1: sucrose binding protein; SBP2: sucrose uptake; transgenic plant;
KW seed: carbohydrate content; soybean.
OS Glycine max.
PN M09853086-A1.
PD 26-NOV-1998.
PF 21-MAY-1998; U10465.
PR 22-MAY-1997; US-047568.
PI (UNIW) UNIV WASHINGTON STATE RES FOUND.
PI Chao WS, Grimes HD:
DR MPI: 99-070155/06.
PT New modified plant sucrose binding proteins - used to develop
PT transgenic plants which can have enhanced or decreased sucrose
PT uptake activity in developing seeds
PS Claim 13b; Page 37-38; 58pp: English.
CC This sequence represents a novel sucrose binding protein, SBP2 isolated
CC from Glycine max. This protein is used in a method resulting in the
CC production of a modified plant sucrose binding protein (SBP) which has a
CC modified amino acid sequence compared to a corresponding wild-type SBP,
CC and where expression of the modified SBP in a yeast assay system confers
CC enhanced sucrose uptake compared to the corresponding wild-type SBP.
CC The products of the invention can be used for producing transgenic plants
CC which have modified sucrose uptake activity, particularly in developing
CC seeds. Enhanced sucrose uptake activity in developing seeds may be
CC desirable where it is an advantage to increase the carbohydrate content
CC of the seed (e.g. where the seed is the primary plant material harvested,
CC such as soybean). In contrast, decreased sucrose uptake activity in
CC seeds might be desirable where the vegetative material of the plant is
CC harvested. The SBP regulatory regions confer specific or enhanced
CC expression in developing seeds and so may be used to express any
CC transgene in developing seeds.
Sequence 489 AA:

[illegible][illegible]

RESULT 9
 ID W90339 standard; protein; 524 AA.
 AC W90339;
 DT 24-MAY-1999 (first entry)
 DE G. max SBP1 protein.
 KW SBP1; sucrose binding protein; SBP2; sucrose uptake; transgenic plant;
 RM seed; carbohydrate content; soybean.
 OS Glycine max.
 PN W069853086-A1.
 PD 26-NOV-1998.
 PF 21-MAY-1998; 010465.
 PR 22-MAY-1997; US-047568.
 PI (UNIW) UNIV WASHINGTON STATE RES FOUND.
 PA Chao WS, Grimes HD;
 PI WPI: 99-070155/06.
 PR New modified plant sucrose binding proteins - used to develop
 PT transgenic plants which can have enhanced or decreased sucrose
 PT uptake activity in developing seeds
 PS Disclosure; Page 34-36; 58pp; English.
 CC This sequence represents a novel sucrose binding protein, SBP1 isolated
 CC from glycine max. This protein is used in a method resulting in the
 CC production of a modified plant sucrose binding protein (SBP) which has a
 CC modified amino acid sequence compared to a corresponding wild-type SBP,
 CC and where expression of the modified SBP in a yeast assay system confers
 CC enhanced sucrose uptake compared to the corresponding wild-type SBP.
 CC The products of the invention can be used for producing transgenic plants
 CC which have modified sucrose uptake activity, particularly in developing
 CC seeds. Enhanced sucrose uptake activity in developing seeds may be
 CC desirable where it is an advantage to increase the carbohydrate content
 CC of the seed (e.g. where the seed is the primary plant material harvested,
 CC such as soybean). In contrast, decreased sucrose uptake activity in
 CC seeds might be desirable where the vegetative material of the plant is
 CC harvested. The SBP regulatory regions confer specific or enhanced
 CC expression in developing seeds and so may be used to express any
 CC transgene in developing seeds.
 CC Sequence 524 AA:

Query Match 20.3%; Score 998; DB 1; length 524;
 Best Local Similarity 37.4%; Pred. No. 3,016-75;
 Matches 167; Conservative 122; Mismatches 137; Indels 20; Gaps 18

Db 49 QQQQYTEDKRV-C-LQ-SCDRHYHMKQREKQIOEFREKKESRREREQQEQHEQD 106
 QY 190 QQRRT-ECRCRRCCQOEERQOYQQRRCRCEQQRHGSGDLINQRGSGRYEGEGRQS 248
 Db 107 ENPYIFEEEDKQFEFRVETEGGRIVLKKETFKSLQGIENFRALIAEAAHFEVSPRH 166
 QY 249 DNPYFDE-RGLSTRFRTEEGHIVLENFGRSKILBALNRYRLVLEAPNAPVLPHTL 307
 Db 167 DSEVFENIKGRAVLGVSESEETKILLEGDMHIEAGPIYVNVNDEMDKLFALMHI 226
 QY 308 DADILVLTGGRGLKMIHRNRETSYMLECGDVIIRIAGTTFYLIINDNRRLLIAF-L 366
 Db 227 PVSYSTPKCEEFPAQGRDSEVLSAFSMNVLOALQITPKLENVFDQONEGSIIRIS 286

```

OY 367 Q-TISTPGQYKEFFPAGGONPEYLSFSTSEILAEALINTQAEIRGLVGOOREGVIISAS 425
DB 287 REOVYALAPTKSSW-WPF--GGES-KPOFNIFSKRPITISNGYGRLETVPPDDDEKSMQ 342
OY 426 QOEIRLELTRDSESRMRHIRRGESSRGPNLFRKRLYKNKYGQAEVPEDEY-RQ-LQ 463
DB 343 RLNLMLTFNITQKSMSTIYNSHATIALYIDGRHLQISCPHMSR-SSH--SK-HDK 398
OY 484 DMVSYFANITQSGMMGPFFNTKSTKYVVVVASGEADYEMACPILSGHGRGGRKHEE 543
DB 399 SSPS-YHRISSDLKPGVNVFVPPCHPEVTIASKNENLMICFEMNADNK-KTFPAGDK- 455
OY 544 EEDVHTEQVAFARLSKRAPIVPGHVPVAVSSGNNELLLEAFGINAQNHENF-LAGRER 602
DB 456 NIVSLLDNVAKELAFNPSEMVNCF 481
OY 603 NVLQIEPQAMELAFAPRKEVELEF 628

RESULT 10
ID W62838 standard; Protein; 605 AA.
AC W62838;
DT 27-OCT-1998 (first entry)
DE Glycine max antimicrobial protein.
KW antimicrobial protein; Infestation; control.
OS Glycine max.
PN W09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (REFR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PT Bremer NT, Goulter KC, Green JL, Manners JM, Marcus JP;
   WPI; 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
   useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 63-65; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
   be used to control microbial infestations in plants and mammalian
   CC animals.
SQ Sequence 605 AA;

Query Match 20.0%; Score 981; DB 1; Length 605;
Best Local Similarity 32.8%; Pred. No. 9,98e-74;
Matches 194; Conservative 160; Mismatches 195; Indels 43; Gaps 32.
DB 38 CLOSENSESDSYRNAQ-CHARCNLLKVEKECEGEIPRPRPOHPR-EPQOPEKEE 95
OY 85 CQRRCQOEGSPRQOQCQRCKEICEEETVNRQNDPQO-YEQOERQRIETPRHM 143
DB 96 DEDEQRPPIFPPOPPOBE--EHOOREOE-WP-R-KEE--KREKSEED-EDDEE 147
OY 144 QYCOQCEKERYEKEKRYEEOQREDEKEERKEEDNRDPOQREYEDCRRCEQ 203
DB 148 ODERO-FPPRPPHOKEREDEDEEORESEES-EDSELRRHKNKNFLEGSRFEE 205
OY 204 QOEPRQOQORCROEORQROGRGDLINPORGSGRYESEE-EKSD-MEYTFDESLST 261
DB 206 LFNQYGRITVLOFNOFQSPOLNLDYRILEFNSKPNITLLPNHADADVLITLNGTAT 265
OY 262 RFTTEGHSIVLENFGRSKLLALKNRYVLEANPNFVLPHTHDADAILVTVGGRKA 321
DB 266 LSLVNNDDRDSTYLGSGDALRVPSTGYTVVNDNNENLRLTLAIPVKNPKGRFESFFLS 325
OY 322 LKMIHRDNRRESYMLECGDVIRIPAGTTFYLINDNNEERLHIAFKLQISPGQKEFFPA 381
DB 326 STPAQOSYLOGEFRNLLEASDYTKFEEINKVLESREEGQOQSGORLOESYIVISKEQIR 385
OY 382 GGNPEPITLSTFKEITLLEALNTQ-AE--R-L--R--GV-L-GQOR--EGVITISQEOQIR 430
DB 386 ALSRAKSSSRK-TI--SSSD-K-PNLRSRDPYISNKLKFEETPEPKNPOLRDLIDFL 440
OY 431 ELT-RDSESRMRHIRRGESSRGPNLFRKRLYKNKYGQAEVPEDEYROLQDMDVAV 489

```

```

DB 441 SIYDNEGALLPHFNASKAIVILYINEGDANIELV-G-LKEOQOEQO--QEOPLEVR- 494
OY 490 FIANTQSGMMGFEFTSTKYVVVVASGEADYEMACPILSGHGRGGRKHEEEDVHX 549
DB 495 -KRALSEODIFIVTAGIPVY-VNAT-S-NLNFPAIGIAENNRFLAGSDNNYISQIP 551
OY 550 EOYKALSRREALVYVGVHVPVAVSSGNNELLLEAFGINAQNHENFLAGRENNVLOQIE 609
DB 552 SOYQELAFPGSAQAVELKLNQRESYFYVDAQPKKKEGNGKRGK--PLSSIL 601
OY 610 PQAMELAFAPRKEVELELNSODESIFFPGRHOQOQSSRSTKQOOPLYSIL 661

RESULT 11
ID W90340 standard; Protein; 444 AA.
AC W90340;
DT 24-MAY-1999 (first entry)
DE G. max truncated SBP1 protein.
KW SBP1; sucrose binding protein; SBP2; sucrose uptake; transgenic plant;
   seed; carbohydrate content; soybean.
OS Glycine max.
PN W09853086-A1.
PD 26-NOV-1998.
PF 21-MAY-1998; U10465.
PR 22-MAY-1997; US-047568.
PA (UNIT ) UNIV WASHINGTON STATE RES FOUND.
PI Chao WS, Grimes HD;
   WPI; 99-070155/06.
PT New modified plant sucrose binding proteins - used to develop
   transgenic plants which can have enhanced or decreased sucrose
   uptake activity in developing seeds
PT Claim 7; Page 36-37; 58pp; English.
PS This sequence represents a novel sucrose binding protein, SBP1 isolated
   from Glycine max. This protein is used in a method resulting in the
   CC production of a modified plant sucrose binding protein (SBP) which has a
   CC modified amino acid sequence compared to a corresponding wild-type SBP,
   CC and where expression of the modified SBP in a yeast assay system confers
   CC enhanced sucrose uptake compared to the corresponding wild-type SBP.
   CC The products of the invention can be used for producing transgenic plants
   CC which have modified sucrose uptake activity, particularly in developing
   CC seeds. Enhanced sucrose uptake activity in developing seeds may be
   CC desirable where it is an advantage to increase the carbohydrate content
   CC of the seed (e.g. where the seed is the primary plant material harvested,
   CC such as soybean). In contrast, decreased sucrose uptake activity in
   CC seeds might be desirable where the vegetative material of the plant is
   CC harvested. The SBP regulatory regions confer specific or enhanced
   CC expression in developing seeds and so may be used to express any
   CC transgene in developing seeds.
SQ Sequence 444 AA;

Query Match 19.0%; Score 933; DB 1; Length 444;
Best Local Similarity 37.8%; Pred. No. 1,94e-69;
Matches 154; Conservative 110; Mismatches 126; Indels 17; Gaps 15;
DB 49 QOQQTGEGDRVC-LQ-SCDRIYHMKQERKQIOETREKKEESREDEEOQNEHOD 106
OY 190 QOREY-EDCRRCEOEQPRQOQCORCROEORQROGRGDLINPORGSGRYESEEKRS 248
DB 107 EMPYIEEKDEETVEDEGRIYVLFKFTESKLLQSTENFRLAILARATYSPNHF 166
OY 249 DNPYFDE-RSLSTRFEEGHSIVLENFYGRSKLLRLKRYVLEANPNFVLPHTL 307
DB 167 DEVEVFENKGRAVGLVSESETEKITLEPGMIHPIPGTPIYVNRBENKLFAMHI 226
OY 308 DADAILVYTGSGALKMHIHRDNRRESYMLECGDVIRIPAGTTFYLINDNNEERLHIAK-L 366
DB 227 PVSYSTPGKFEFFAPGRDPESVLSAFSWNYLQAALQTPKGLTENVEDQOQNEGSIFRIS 286
OY 367 Q-TISTPGQYKEFFPAGGONPEYLSFSTSEILAEALINTQAEIRGLVGOOREGVIISAS 425
DB 287 REOVYALAPTKSSW-WPF--GGES-KPOFNIFSKRPITISNGYGRLETVPPDDDEKSMQ 342
OY 426 QOEIRLELTRDSESRMRHIRRGESSRGPNLFRKRLYKNKYGQAEVPEDEY-RQ-LQ 483

```

ID	Accession	Protein	Score	DB	Length
Db	343	RLNMLTFTNTTORSMSTIHNSATKIALYIDRGHLQICPMHSSR--SK-HDK	398		
Qy	484	DMDVSVFIAINTIGSGMGMPFFNTSTRTKVVVASGADVEAMCAPHLSGRHGGRGKREE	543		
Db	399	SSPS-YHRISSDLRKGVVPPGPFVATISNKNENLMICEFNAR	444		
Qy	544	EDVHYEYKARLSKREAIIVPVGHVVFVSSGNENLLFAFGINAQ	590		
RESULT 12					
ID	W22150	standard; protein; 626 AA.			
AC	W22150.1				
DT	29-DEC-1997	(first entry)			
DE		peanut allergen Ara hi.			
KW		peanut; seed storage protein; allergen; allergy; hypersensitivity;			
KW		vaccine; anaphylactic shock; immunotherapy; therapy;			
OS		monoclonal antibody; ELISA; analysis; Ara hi.			
OS		Arachis hypogaea strain Florunner.			
PH	Key	Location/Qualifiers			
FT	Peptide	1..22			
FT		/label= Sig-peptide			
FT	Protein	23..626			
FT		/label= Mat-protein			
FT	Modified_site	521..523			
FT		/note= "N-glycosylation site"			
PN	MO9724139-A1.				
PD	10-JUL-1997.				
PR	23-SEP-1996.	U15222.			
PR	04-MAR-1996.	US-610424.			
PR	29-DEC-1995.	US-009455.			
PA	(U0VAR-) UNIV ARKANSAS.				
PI	Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;				
DR	WPI: 97-363453/33.				
DR	N-PSDB: T76613.				
PT		peanut allergens Ara hi and Ara hII - used for vaccination and in			
PT		two-site monoclonal antibody based ELISA			
PS	Claim 31; Page 172; 354pp; English.				
CC		This polypeptide comprises major peanut allergen Ara hi (W22149).			
CC		Its sequence was deduced from cDNA clone P41b (T76613), isolated			
CC		from peanut seed cDNA using a primer (see T76616) based on an			
CC		isolated Ara hi peptide (see W24206). The sequence shows			
CC		significant homology with the vicilin family of seed storage			
CC		proteins of other legumes. The allergen is recognised by serum			
CC		IGE from a large proportion of individuals with peanut			
CC		hypersensitivity. Ara hi and Ara hII (see W24164) can be used to			
CC		raise monoclonal antibodies which are used in a specific two-site			
CC		MAb ELISA for the detection of Ara hi or Ara hII (claimed). IGE-			
CC		binding Ara hi antigen epitopes (see W24165-87) may be used in			
CC		vaccines to protect against allergic reactions to peanut allergens,			
CC		e.g. anaphylactic shock.			
CC		Sequence 626 AA;			
Query Match 19 0%: Score 933; DB 1; Length 626;					
Best Local Similarity 33.6%: Fred. No. 1.94e-69;					
Matches 195; Conservative 156; Mismatches 188; Indels 42; Gaps 30;					
Db	38	CAORCLASGQEPDLKQKACESRCTK-LEVDPCVYDPRGHTGTNORSPG-ERTGR	95		
Qy	102	CQRKKEKCEEEFEENKRRDPQOYEOQCEQCHNTEPR-HMOTCOORERTREKERRK	160		
Db	96	QPGDYDDRRQPRREBGR-WGAPAPRE-REFE-ED-WRQ-PREDWRPSPH-QQP-R-KI	147		
Qy	161	QOKRVEEQREDEKYEKMEEDNKRPQREVEDRCRRQEQEPQOYOCQQRCEQ	220		
Db	148	RPEGEFG--QENGTPQSHVR-BETSANPFYPPSRFSTRIGYNQNGRIRVLOREDQRS	203		
Qy	221	ROHGGGGLINPORGSGSRVEEGEKOSDNFYEDERSLSTRFTEEGHISVLENFGRS	280		
Db	204	ROFOQLNHRIVLOEARKNTLVLRKHAADNIIIVQOGQAVTVYANNNKRSFNLDEGHA	263		
Qy	281	KLRRLKNTVRLVLEANNNAFLPLTHADADIIIVTGGRGALKIHHNDNRSYVLEGGDV	340		

```

Db      264  LRLPSGSLSYLINRDRDNQNLVAKTSMVYVQPGQEDFEFPASSSDQSSYLDOGFANLTLEA 323
Qy      341  IRIPIPGTFYLLINRNNRNLRLAKFLQITSPGQYKEFFPAGGONPEYLYSTFEKLELLA 400
Db      324  AFNAFENLRRLLEENNGGEDEEGRGRRMSTRSENNEGVYKVRSKREVEELTKHAKSV 383
Qy      401  ALNTQAEKRLGVL-----GQORE-GVLI-SA--SQE-Q--IRELTRPD-SE-SRR-WHI 444
Db      384  SKKSGSEEDGITNPILNREGEEDPDSNNFGKLEFYKPKDKNPOLODLMDMLTCVEIKEGAL 443
Qy      445  -RRGE--SS-RGPYNLFLNRKRLPYLSNKYGAQYEVKPED-YRQLODMQVSVFIANTQCSM 499
Db      444  MLPHNRSKAMTVVYVYVNGKTGNLELVAVKEDQOQGRREDEDEDEEESGRREVRRTIAR 503
Qy      500  MGPFNTSTRKTVVVAASEADVEMACPHLSRHGGRGRGRHE--EEDVHYEQVK---R 555
Db      504  LKEGVFTMPAAHPVA-INASSE-LHLFGFINAENNRHRIADCKDVIDQIEKQAKDL 561
Qy      556  LSKRAITVYVPGHPVPEYSSGNNELTFAPFIMQNNHENTLAGRERNVLDQIEPQAMEL 615
Qy      616  AFAARKEVEELFNDSODESIFFPGRHQOQSSSTKQOCP 656

RESULT 13
ID      13  W22149 standard; Protein;#614 AA.
AC      W22149;
MT      29-DEC-1997 (first entry)
DE      Peanut allergen Ara hi.
KW      Peanut; seed storage protein; allergen; allergy; hypersensitivity;
        vaccine; anaphylactic shock; immunotherapy; therapy;
        monoclonal antibody; ELISA; analysis; Ara hi.
OS      Arachis hypogaea strain Florunner.
FH      Key
FT      Peptide
        1..22
        /label= Sig-peptide
FT      Protein
        23..614
        /label= Mat-protein
FT      Modified_site
        521..523
        /note= "N-glycosylation site"

PN      W09724139-A1.
PD      10-JUL-1997.
PR      23-SEP-1986; U15222
        04-MAR-1986; US-610424.
        29-DEC-1995; US-009455.
PA      (UYAR-) UNIV ARKANSAS.
        Bannon GA, Burks AW, Cockrell G, Heim RW, Stanley JS;
        WPI; 97-363453/33.
DR      N-PSDB: T76612.
PT      Peanut allergens Ara hi and Ara hII - used for vaccination and in
        two-site monoclonal antibody based ELISA
        Claim 31; Page 169; 354pp; English.
FC      This polypeptide comprises major peanut allergen Ara hi (W22149).
        Its sequence was deduced from cDNA clone P17 (T76612), isolated
        from peanut seed cDNA using a primer (see T76616) based on an
        isolated Ara hi peptide (see W24206). The sequence shows
        significant homology with the vicilin family of seed storage
        proteins of other legumes. The allergen is recognised by serum
        IGE from a large proportion of individuals with peanut
        hypersensitivity. Ara hi and Ara hII (see W24164) can be used to
        raise monoclonal antibodies which are used in a specific two-site
        Mab ELISA for the detection of Ara hi or Ara hII (claimed). IGE-
        binding Ara hi antigen epitopes (see W24165-87) may be used in
        vaccines to protect against allergic reactions to peanut allergens,
        e.g. anaphylactic shock.
SQ      Sequence 614 AA;

Query Match 18.8%; Score 922; DB 1; Length 614;
Best Local Similarity 34.0%; Pred. No. 1,86e-66;
Matches 192; Conservative 158; Mismatches 173; Indels 42; Gaps 30;

```



```

QY 119 QRDQOQYEQ-CQERCCQHETEPHMQTCQRCERRERKQCKRKEEQQRDEKYE 177
DB 101 PR-REEGRMGAPAPRE-RE-REE-DMROPREDWR--RPSHQPRKRIPEEGREGEOEMGT 154
QY 178 ERKMEEDNKRBP-QQREYEDCRRCEDQEPHQYOCQRCRCEQORRGDGLINPORG 236
DB 155 PGS-EVBEETSRRNPFFPSRRFSTRYGONGRIRVLOREDFORSKOFONLONHRIVOIEA 213
QY 237 SGREVEEGEEKQSDNPYYFDERSLSTRFTEEGHISYLENFGRSKLLRALKNVRLVLEA 296
DB 214 RPNVLVLPKHADNLIIVIOGOATVYANGNNKSFNLDEGHALRIPSGFISYILNRHD 273
QY 297 NPNAFVLPHTLDADALLVYGGGALKMIRHDSREYNLECGDVIRIPAGTFFYLNRDN 356
DB 274 NONLRVAKISMPTVPGQEDFEPASSRDOSSYLQGSFNTLEAFNAFNEIRVLEE 333
QY 357 NERHIKFLQITISTPGQYKEFFPAGGONPEPILSTFSKELFAALNTQAEIRGLV--- 413
DB 334 MAGDEEERGQRRSTRSSDNEGVIYVSKHEVOELTKHAKSVSKGSEEDITNPINLR 393
QY 414 --GQORE-GVII-SA-SQE-Q--IRELTRDS-E-SRR-WHI-RGGES-S-RGPYNLF 458
DB 394 DGEEDLSNNGRLEFVYVDPKKNPOLDDMLTCVEITKEGALMLPHNSKAMVIVYVNGK 453
QY 459 NKRPILSNKYGQAEVYKPED-YRQLODMVSVFIANTQGSMMGPFFNTRSTKVYVVASG 517
DB 454 TGNLELVAVRKEQOGRREQEWEDEEEDSGSREVRRTARLKEGDFVIMPAHPV 513
QY 518 EADVEACPHLSGHRGGRGKRHEEDVHYE---QVK---ARLSKREAIYVPGHPV 570
DB 514 A-1NASSE-LHLFGFINAENNRIFLAGDKNDVIDQIEKQAKDLAFPGSGEOVEKILNK 571
QY 571 VFVSSGNNELLFAFGINQNNHENFLAGREVRVLIQIIEPQAMELAAPRKEVEELFNS 630
DB 572 QRES-HFVSARPOSQSPSSPEKEDQ 595
QY 631 QDESIFPPGRHQOQSSRSTKQOQ 655

```

RESULT 14

```

ID W62834 standard; Peptide: 614 AA.
AC W62834;
DT 27-OCT-1998 (first entry)
DE Arachis hypogaea antimicrobial protein.
KM antimicrobial protein; Infestation; control.
OS Arachis hypogaea.
PN W09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NJ, Goulter KC, Green JL, Manners JM, Marcus JP;
WPI: 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 55-57; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 614 AA:

```

Query Match 18.8%; Score 922; DB 1; Length 614;

Best Local Similarity 34.0%; Pred. No. 1,86e-68;

Matches 192; Conservative 158; Mismatches 173; Indels 42; Gaps 30;

```

DB 45 QQEPDILKQKACSRCTKLEYDPCVDTGAT-NQHPPEBRTG-R-QFGYDDDR-Q 100
QY 119 QRPQOQYEQ-CQERCCQHETEPHMQTCQRCERRERKQCKRKEEQQRDEKYE 177
DB 101 PR-REEGRMGAPAPRE-RE-REE-DMROPREDWR--RPSHQPRKRIPEEGREGEOEMGT 154
QY 178 ERKMEEDNKRBP-QQREYEDCRRCEDQEPHQYOCQRCRCEQORRGDGLINPORG 236

```

```

DB 155 PGS-EVBEETSRRNPFFPSRRFSTRYGONGRIRVLOREDFORSKOFONLONHRIVOIEA 213
QY 237 SGREVEEGEEKQSDNPYYFDERSLSTRFTEEGHISYLENFGRSKLLRALKNVRLVLEA 296
DB 214 RPNVLVLPKHADNLIIVIOGOATVYANGNNKSFNLDEGHALRIPSGFISYILNRHD 273
QY 297 NPNAFVLPHTLDADALLVYGGGALKMIRHDSREYNLECGDVIRIPAGTFFYLNRDN 356
DB 274 NONLRVAKISMPTVPGQEDFEPASSRDOSSYLQGSFNTLEAFNAFNEIRVLEE 333
QY 357 NERHIKFLQITISTPGQYKEFFPAGGONPEPILSTFSKELFAALNTQAEIRGLV--- 413
DB 334 MAGDEEERGQRRSTRSSDNEGVIYVSKHEVOELTKHAKSVSKGSEEDITNPINLR 393
QY 414 --GQORE-GVII-SA-SQE-Q--IRELTRDS-E-SRR-WHI-RGGES-S-RGPYNLF 458
DB 394 DGEEDLSNNGRLEFVYVDPKKNPOLDDMLTCVEITKEGALMLPHNSKAMVIVYVNGK 453
QY 459 NKRPILSNKYGQAEVYKPED-YRQLODMVSVFIANTQGSMMGPFFNTRSTKVYVVASG 517
DB 454 TGNLELVAVRKEQOGRREQEWEDEEEDSGSREVRRTARLKEGDFVIMPAHPV 513
QY 518 EADVEACPHLSGHRGGRGKRHEEDVHYE---QVK---ARLSKREAIYVPGHPV 570
DB 514 A-1NASSE-LHLFGFINAENNRIFLAGDKNDVIDQIEKQAKDLAFPGSGEOVEKILNK 571
QY 571 VFVSSGNNELLFAFGINQNNHENFLAGREVRVLIQIIEPQAMELAAPRKEVEELFNS 630
DB 572 QRES-HFVSARPOSQSPSSPEKEDQ 595
QY 631 QDESIFPPGRHQOQSSRSTKQOQ 655

```

RESULT 15

```

ID W90342 standard; protein: 409 AA.
AC W90342;
DT 24-MAY-1999 (first entry)
DE G.. max truncated SBP2 protein.
KM SBP1; sucrose binding protein; SBP2; sucrose uptake; transgenic plant;
KM seed; carbohydrate content; soybean.
OS Glycine max.
PN W09853086-A1.
PD 26-NOV-1998.
PF 21-MAY-1998; U10465.
PR 22-MAY-1997; US-047568.
PA (UNTW ) UNIV WASHINGTON STATE RES FOUND.
PI Chao WS, Grimes HD;
WPI: 99-070153/06.
PT New modified plant sucrose binding proteins - used to develop
PT transgenic plants which can have enhanced or decreased sucrose
PT uptake activity in developing seeds
PS Claim 7; Page 39-40; 58pp; English.

```

This sequence represents a novel sucrose binding protein, SBP2 isolated from Glycine max. This protein is used in a method resulting in the production of a modified plant sucrose binding protein (SBP) which has a modified amino acid sequence compared to a corresponding wild-type SBP, and where expression of the modified SBP in a yeast assay system confers enhanced sucrose uptake compared to the corresponding wild-type SBP. The products of the invention can be used for producing transgenic plants which have modified sucrose uptake activity, particularly in developing seeds. Enhanced sucrose uptake activity in developing seeds may be desirable where it is an advantage to increase the carbohydrate content of the seed (e.g. where the seed is the primary plant material harvested, such as soybean). In contrast, decreased sucrose uptake activity in seeds might be desirable where the vegetative material of the plant is harvested. The SBP regulatory regions confer specific or enhanced expression in developing seeds and so may be used to express any transgene in developing seeds.

SQ Sequence 409 AA:

Query Match 18.1%; Score 890; DB 1; Length 409;

Best Local Similarity 37.8%; Pred. No. 1.33e-65;

Matches 146; Conservative 105; Mismatches 117; Indels 18; Gaps 14;

```

Db      35 EDPETVTCCKHOCQOO--ROYTESDKRTCLQOCDSMKOEKREKOVEEETREKEEHOEHE 92
      190 QGREYEDCRRCEQOEPRQOYCQRR-CREQ-QR-QHGRGDLINPQRGSGRYEEG--EE 245
Db      93 EEDENPYVEEDKDFSTRVETEGGSIRVLKKEFEKSKLLQGIENFRLAILEARAHFVSP 152
      246 KQSDNPYFDE-RSLSTRPTEEGHISVLENFYGRSKLLRALKNYRLVYLEANPNAFVLP 304
Db      153 RHFDSEVYLFNKGRAVLGLVRESETEKITLEPQDMIHIPAGTPLYIVRDENEKILLAM 212
      305 THLDADAILVTGGRGALKMHIRDNRESYNLECGDVIRIPAGTFFYLINRDNNERLHIAK 364
Db      213 LHIPTVTPGKEFEFFGPGGDESVLSAFSMNVLOALQTPKGLERLEFNQONEGSIFKI 272
      365 FLQITISTPGQYKEFFPAGQNPPEYLISTFSKILEALNTQAERLQVILGQOREGYIISA 424
Db      273 SREVRALAPTKSSW-WPF--GSES-KAQENIFSKRPTFSNGYRLTEVGPDEKSWLQ 328
      425 SQEQIRELTRDSESRRWHIRRGESSRGPYNLFNKRPLYSNKYQAYEVKPEDYRQ-LQ 483
Db      329 RUNMLTFTNITQRMSTIHYNSHATKIALVMDGRGHQISCPHMSR--S--DSK-HDK 383
      484 DMDVSVFIANITQSGSMGPFENRSTKVVYVASGEADVEMACPHLSGRHGRGGRKHEE 543
Db      384 SSPS-YHRISADLKPGWVFPVPGHP 408
      544 EEDVHYEQVKARLSKREAIYVPVGH 569
Qy

```

Search completed: Sat May 13 08:09:07 2000
 Job time : 29 secs.

